Algorithmes d'optimisation

5 février 2019

Table des matières

Introduction

Motivation Dans de nombreuses applications, la formulation naturelle du problème qu'on cherche à résoudre est un problème d'optimisation :

— Dans la méthode des moindres carrés, on remplace un système linéaire $Ax = b$ surdéterminé et/ou n'ayant pas de solution (par exemple car certaines des égalités se contredisent) par le problème d'optimisation suivant :

$$\min_{x \in \mathbb{R}^N} \|Ax - b\|^2 . \tag{1}$$

Le minimiseur x^* de ce problème vérifie "au mieux" la famille d'équation $Ax = b$. Au contraire, lorsqu'un système linéaire $Ax = b$ admet plusieurs solutions, on peut en sélectionner une en considérant le problème

$$\min_{x \in K} \|x\|^2 \quad K = \{x \in \mathbb{R}^N \mid Ax = b\} \tag{2}$$

— Si $\bar{x} \in \mathbb{R}^N$ représente un signal 1D échantilloné avec du bruit ($\bar{x}_i$ représentant par exemple la mesure effectuée en un temps t_i), on peut débruiter le signal en considérant le problème d'optimisation suivant, où $\lambda > 0$ est un paramètre :

$$\min_{x \in \mathbb{R}^N} \|x - \bar{x}\|^2 + \lambda \sum_{1 \le i \le N-1} |x_{i+1} - x_i|^2 \tag{3}$$

un compromis entre deux comportements : x^* doit être proche de $\bar{x}$ (c'est le rôle du premier terme $\|x - \bar{x}\|^2$ de la fonction optimisée) mais doit également être "régulier", au sens où deux valeurs successives x_i et x_{i+1} doivent être proches (second terme $\sum_i |x_{i+1} - x_i|^2$).

— En finance, on peut considérer le problème de l'optimisation de portefeuille. Étant donnés N actifs, il s'agit de déterminer le pourcentage $x_i \geq 0$ du portefeuille que l'on investit dans l'actif i. Comme on souhaite investir 100% du portefeuille, ce problème d'optimisation est accompagné d'une contrainte $\sum_{1 \le i \le N} x_i = 1$. On pourra donc considérer des problèmes d'optimisation *avec contraintes* de la forme

$$\min_{x \in \Delta} f(x) \text{ où } \Delta = \{x \in \mathbb{R}^N \mid \forall i, x_i \geq 0 \text{ et } \sum_i x_i = 1\}. \tag{4}$$

La fonction f est typiquement de la forme $f(x) = \frac{1}{\varepsilon} |\langle c|x\rangle - r|^2 + \langle x|Qx\rangle$: $c \in \mathbb{R}^N$ représente le rendement des actifs et le premier terme cherche à fixer

le niveau de rendement $\langle c|x\rangle = \sum_i c_i x_i$ à r. Le second terme de la fonction $\langle x|Qx\rangle$ est une mesure de risque : Q est une matrice symétrique mesurant les corrélations entre actifs, et on cherche un investissement minimisant cette corrélation.
— En apprentissage automatique (*machine learning*), de nombreux problèmes peuvent être formulés comme des problèmes d'optimisation. Nous verrons par exemple des problèmes de classification, que l'on résoudra par régression logistique ou par machine à vecteurs support (*support vector machine*).

Pour plus d'exemple, on renvoie au livre de Boyd et Vanderberghe, qui est disponible gratuitement (en anglais) en ligne : `http://web.stanford.edu/~boyd/cvxbook/`.

Problèmes avec/sans contrainte On peut séparer les problèmes en deux grandes classes. Il y a d'une part les problèmes d'optimisation sans contraintes, où l'on cherche à minimiser une fonctionnelle sur $\mathbb{R}^d$ ou sur son domaine de définition *ouvert* (les problèmes (1), (3) et la régression logistique sont de ce type). D'autre part, les problèmes d'optimisation avec contraintes, où l'on cherche à minimiser sur l'ensemble des points de $\mathbb{R}^N$ vérifiant un certain nombre de contraintes d'égalité ou d'inégalité (les problèmes (2), (4) et les machines à vecteurs support sont de ce type).

Convexité Tout les algorithmes et exemples présentés dans ce cours relèvent de l'optimisation *convexe*, où aussi bien la fonction optimisée que le domaine d'optimisation sont supposés convexes. La raison fondamentale pour laquelle on se restreint à ce cas est que pour les problèmes d'optimisation convexe, un minimiseur local est *automatiquement* minimiseur global.

Chapitre 1

Existence, unicité, convexité

Contents

Dans cette première partie, on s'intéresse à un problème de minimisation d'une fonction $f \in \mathcal{C}^1(\Omega)$ où Ω est un ouvert de $\mathbb{R}^d$:

$$\inf_{x \in \Omega} f(x) \tag{P}$$

Définition 1. On appelle :

(i) *infimum* de (P) la valeur $\inf_\Omega f$.

(ii) *minimiseur global* (ou simplement minimiseur) de (P) tout élément $x^* \in \Omega$ vérifiant $f(x^*) = \inf_\Omega f$. On note $\arg\min_\Omega f$ l'ensemble des minimiseurs de f (qui peut être vide).

(iii) On appelle *suite minimisante* pour (P) toute suite $x^{(0)}, \dots, x^{(k)}, \dots$ d'éléments de Ω telle que $\lim_{k\to+\infty} f(x^{(k)}) = \inf_{x\in\Omega} f(x)$.

Remarque 1. Il est possible que le problème (P) n'admette pas de minimiseur : penser par exemple à $f(x) = \exp(x)$ sur $\Omega = \mathbb{R}$, ou $f(x) = x^2$ sur $\Omega =]0, +\infty[$.

Lemme 1. *Il existe une suite minimisante pour le problème* (P).

Démonstration. Par définition de l'infimum, pour tout $k > 0$, il existe un élément $x^{(k)} \in \Omega$ tel que $\inf_\Omega f \leq f(x^{(k)}) \leq \inf_\Omega f + \frac{1}{k}$, soit $\lim_{k\to+\infty} f(x^{(k)}) = \inf_\Omega f$. $\square$

1.1 Existence

Proposition 2. *Soit* $\Omega \subseteq \mathbb{R}^d$ *ouvert et* $f \in \mathcal{C}^0(\Omega)$. *On suppose de plus qu'il existe* $x_0 \in \Omega$ *tel que le sous-niveau* $S = \{x \in \Omega \mid f(x) \leq f(x_0)\}$ *est compact. Alors le problème d'optimisation* (P) *admet un minimiseur global.*

Démonstration. Si $f(x_0) = \inf_\Omega f$ on a déjà l'existence d'un minimum, à savoir le point x_0 lui-même. On suppose donc maintenant que $f(x_0) > \inf_\Omega f$. Soit $(x^{(k)})_{k\geq 0}$ une suite minimisante, qui vérifie donc $\lim_{k\to+\infty} f(x^{(k)}) = \inf_\Omega f < f(x_0)$. Alors, pour k suffisamment grand, on a $f(x^{(k)}) \leq f(x_0)$, soit $x^{(k)} \in S$. Comme l'ensemble S est compact, on peut extraire une sous-suite $(x^{(\sigma(k))})_{k\geq 0}$ qui converge vers un point $x^\infty \in S \subseteq \Omega$. Alors, par continuité de f et par définition d'une suite minimisante on a $f(x^\infty) = \lim_{k\to+\infty} f(x^{(\sigma(k))}) = \inf_\Omega f$, et x^∞ minimise donc f sur Ω. □

Corollaire 3. *Soit* $f \in \mathcal{C}^0(\Omega)$ *où* $\Omega = \mathbb{R}^d$ *et vérifiant la propriété suivante :*

$$\forall x \in \mathbb{R}^d, f(x) \geq C\,\|x\|^p + D,$$

où $C > 0, D \in \mathbb{R}$ *et* $p > 0$. *Alors le problème* (P) *admet un minimiseur.*

Démonstration. Soit $x_0 \in \mathbb{R}^d$ quelconque et $S = \{x \in \mathbb{R}^d \mid f(x) \leq f(x_0)\}$. Pour pouvoir appliquer le théorème précédent, il suffit de démontrer que S est compact. Comme on sait déjà que S est fermé comme sous-niveau d'une fonction continue, il suffit de démontrer que cet ensemble est borné. Or, pour tout $x \in S$, on a

$$C\,\|x\|^p + D \leq f(x) \leq f(x^{(0)})$$

soit $\|x\|^p \leq E := \left|f(x^{(0)}) - D\right|/C$. Ainsi, l'ensemble S est contenu dans la boule centrée en 0 et de rayon $\sqrt[p]{E}$ et est donc borné. □

Corollaire 4. *Soit* $f \in \mathcal{C}^0(\Omega)$ *où* $\Omega = \mathbb{R}^d$ *vérifiant la propriété suivante*[1] *: pour toute suite* $(x^{(k)})_{k\geq 0}$ *vérifiant* $\lim_{k\to+\infty} \left\|x^{(k)}\right\| = +\infty$ *alors* $\lim_{k\to+\infty} f(x^{(k)}) = +\infty$. *Alors* (P) *admet un minimiseur global.*

Démonstration. Soit $x_0 \in \mathbb{R}^d$ quelconque et $S = \{x \in \mathbb{R}^d \mid f(x) \leq f(x_0)\}$. Pour montrer l'existence d'un minimiseur, il suffit de démontrer que S est compact. Par l'absurde supposons S non compact. Comme S est fermé (sous-niveau d'une fonction continue) on en déduit qu'il n'est pas borné. Il existe donc une suite $(x^{(k)})_{k\geq 0}$ d'éléments de S (c'est-à-dire tels que $f(x^{(k)}) \leq f(x_0)$) telle que $\lim_{k\to+\infty} \left\|x^{(k)}\right\| = +\infty$. Ceci contredit note hypothèse. □

1. qu'on peut résumer en disant que “f tend vers l'infini à l'infini”

1.2 Condition nécessaire d'optimalité

Définition 2. Soit $f \in \mathcal{C}^1(\Omega)$, où $\Omega \subseteq \mathbb{R}^d$ est un ouvert. On appelle *gradient* de f en $x_0 \in \Omega$ le vecteur des dérivées partielles, où l'on a noté $(e_i)_{1\leq i\leq d}$ la base canonique de $\mathbb{R}^d$:

$$\nabla f(x) = \left(\frac{\partial f}{\partial e_i}(x)\right)_{1\leq i\leq d} \quad \text{où} \quad \frac{\partial f}{\partial e_i}(x) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(f(x+\varepsilon e_i) - f(x)).$$

Remarque 2 (Calcul du gradient). Soit $f \in \mathcal{C}^0(\Omega)$ et $x \in \Omega$. On rappelle que si l'on peut écrire un développement limité d'ordre 1 pour f en x, de la forme

$$f(x+v) = f(x) + \langle g|v\rangle + \mathrm{o}(\|v\|), \tag{1.1}$$

alors f est différentiable au point x et $\nabla f(x) = g$.

Exemple 1. Considérons $f(x) = \|x\|^2$ sur $\mathbb{R}^d$. En développant le carré de la norme, on obtient $f(x+v) = \|x\|^2 + \langle 2x|v\rangle + \|v\|^2$, qui est de le forme (1.1) avec $g = 2x$. On en déduit que $\nabla f(x) = 2x$, ce qui est conforme avec le calcul.

Théorème 5 (Fermat)**.** *Soit $f \in \mathcal{C}^1(\Omega)$, où $\Omega \subseteq \mathbb{R}^d$ est un ouvert, et x^* un minimiseur de* (P)*. Alors $\nabla f(x^*) = 0$.*

Remarque 3. La contraposée est fausse : prendre $f(x) = x^3$ sur $\Omega = \mathbb{R}$: le point 0 vérifie $f'(0) = 0$ mais n'est pas un minimiseur (même local).

Démonstration. Comme Ω est ouvert, il existe $\varepsilon_0 > 0$ tel que $\mathrm{B}(x^*, \varepsilon_0) \subseteq \Omega$). Soit $1 \leq i \leq d$ et $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, de sorte que le point $x^* + \varepsilon e_i$ appartient à Ω. Comme x^* est un minimiseur de f sur Ω on a $f(x^* + \varepsilon e_i) \geq f(x^*)$. Ainsi,

$$\forall 0 < \varepsilon \leq \varepsilon_0, \quad \frac{f(x^* + \varepsilon e_i) - f(x^*)}{\varepsilon} \geq 0.$$

En passant à la limite, on obtient $\frac{\partial f}{\partial e_i}(x^*) \geq 0$. De même, en considérant le cas $\varepsilon < 0$

$$\forall -\varepsilon_0 \leq \varepsilon < 0, \quad \frac{f(x^* + \varepsilon e_i) - f(x^*)}{\varepsilon} \leq 0,$$

d'où l'on tire en passant à la limite $\frac{\partial f}{\partial e_i}(x^*) \leq 0$, soit *in fine* $\frac{\partial f}{\partial e_i}(x^*) = 0$. □

1.3 Convexité et condition suffisante d'optimalité

Définition 3. Un sous-ensemble X de $\mathbb{R}^d$ est dit *convexe* s'il contient tout segment reliant deux de ses points, c'est-à-dire :

$$\forall (x, y) \in X, \forall t \in [0, 1], \ (1-t)x + ty \in X.$$

Exemple 2. (a) Pour tout $x \in \mathbb{R}^d$ et $r \geq 0$, la boule centrée en x et de rayon r (ouverte ou fermée) est convexe.

(b) Pour tout vecteur $v \in \mathbb{R}^d$ et tout $b \in \mathbb{R}$, l'ensemble $\{x \in \mathbb{R}^d \mid \langle x|v\rangle \leq b\}$ est un ensemble convexe, appelé *demi-espace.*

(c) L'intersection $\cap_{i\in I} X_i$ d'une famille quelconque $(X_i)_{i\in I}$ de convexes est convexe.

Définition 4. Soit $X \subseteq \mathbb{R}^d$ convexe. Une fonction $f : X \to \mathbb{R}$ est *convexe* si

$$\forall (x, y) \in X, \forall t \in [0, 1], \;\; f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

Exercice 1. Soit $f : \Omega \to \mathbb{R}$ une fonction convexe. Montrer que l'ensemble

$$\{x \in \Omega \mid f(x) \leq C\}$$

est convexe quel que soit C. En déduire que l'ensemble des minimiseurs de (P) est un ensemble convexe fermé (possiblement vide).

Proposition 6. *Soit $f \in \mathcal{C}^1(\Omega)$, $\Omega \subseteq \mathbb{R}^d$ ouvert et $X \subseteq \Omega$ convexe. Alors les propositions suivantes sont équivalentes :*

(i) f est convexe sur X,

(ii) $\forall x, y \in X$, la fonction $g : t \in [0, 1] \mapsto f((1-t)x + ty)$ est convexe.

(iii) $\forall x, y \in X, f(y) \geq f(x) + \langle y - x|\nabla f(x)\rangle$,

(iv) $\forall x, y \in X, \langle \nabla f(x) - \nabla f(y)|x - y\rangle \geq 0$.

Lemme 7. *Soit $X \subseteq \mathbb{R}^d$ convexe et $f \in \mathcal{C}^1(\Omega)$. Étant donnés $x \in \Omega$ et $v \in \mathbb{R}^d$, on définit $x_t = x + tv$ et $g(t) = f(x_t)$. Alors,*

$$g'(t) = \langle \nabla f(x_t)|v\rangle. \tag{1.2}$$

Démonstration de la proposition 6. (i)$\Longleftrightarrow$(ii) conséquence directe de la définition.

(ii)$\Longrightarrow$(iii). Soit x, y dans X. Si f est convexe, alors la fonction $g : t \mapsto f(x_t)$ où $x_t = (1-t)x + tv = x + t(y - x)$ l'est aussi. Ainsi,

$$g(t) = g((1-t)0 + t1) \leq (1-t)g(0) + tg(1),$$

soit $g(1) \geq g(0) + \frac{1}{t}(g(t) - g(0))$ pour $0 < t \leq 1$. En passant à la limite lorsque $t \to 0$, on obtient $g(1) \geq g(0) + g'(0)$, soit par le lemme, $f(x) + \langle \nabla f(x)|y - x\rangle + \leq f(y)$.

(iii)$\Longrightarrow$(ii). Soit $x, y \in X$, $t \in [0, 1]$ et $z = (1-t)x + ty$. Alors, l'inégalité (ii) appliquée à z, x puis z, y donne

$$f(x) \geq f(z) + \langle x - z|\nabla f(z)\rangle = f(z) + t\langle x - y|\nabla f(z)\rangle$$

$$f(y) \geq f(z) + \langle y - z|\nabla f(z)\rangle = f(z) + (1-t)\langle y - x|\nabla f(z)\rangle$$

La somme de $(1-t)$ fois la première inégalité et t fois la seconde donne

$$(1-t)f(x) + tf(y) \geq f(z).$$

(iii)$\Longrightarrow$(iv) Il suffit de sommer l'inégalité (iii) et la même inégalité où l'on a inversé le rôle de x et y.
(iv)$\Longrightarrow$(iii) Soit encore $g : t \mapsto f(x_t)$ où $x_t = (1-t)x + tv = x + t(y-x)$. Comme $g'(t) = \langle \nabla f(x_t) | y - x \rangle$, (lemme 7) l'inégalité (iv) appliquée en x_s et x_t (où $t > s$) nous donne

$$g'(t) - g'(s) = \langle \nabla f(x_t) - \nabla f(x_s) | y - x \rangle = \frac{1}{t-s} \langle \nabla f(x_t) - \nabla f(x_s) | x_t - x_s \rangle \geq 0,$$

et g' est donc croissante sur $[0,1]$. Ainsi, $g(1) = g(0) + \int_0^1 g'(t) \mathrm{d}t \geq g(0) + g'(0)$, ce qui se traduit par $f(y) \geq f(x) + \langle y - x | \nabla f(x) \rangle$ comme souhaité. □

Théorème 8. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe et $f \in \mathcal{C}^1(\Omega)$ convexe. Alors $x^* \in \Omega$ est un minimiseur de* (P) *si et seulement si $\nabla f(x^*) = 0$.*

Démonstration. Le théorème de Fermat nous donne déjà le sens direct. Pour la réciproque, il suffit de remarquer que si $\nabla f(x^*) = 0$, la proposition précédente donne

$$\forall y \in \Omega, f(y) \geq f(x^*) + \langle y - x^* | \nabla f(x^*) \rangle = f(x^*),$$

de sorte que x^* est bien un minimiseur de (P). □

1.4 Stricte convexité et unicité du minimiseur

Définition 5. Soit $X \subseteq \mathbb{R}^d$ un ensemble convexe. Une fonction $f : X \to \mathbb{R}$ est strictement convexe si

$$\forall x \neq y \in X, \forall t \in]0,1[, \;\; f((1-t)x + ty) < (1-t)f(x) + tf(y).$$

Proposition 9. *Soit $X \subseteq \mathbb{R}^d$ convexe et $f : X \to \mathbb{R}$ strictement convexe. Alors f admet au plus un minimiseur sur X.*

Démonstration. Par l'absurde, supposons que f admette deux minimiseurs distincts $x^* \neq y^* \in \Omega$. Par convexité de Ω, le point $z^* = \frac{1}{2}(x^* + y^*)$ appartient à Ω. Alors, par stricte convexité de la fonction f on a $f(z^*) < \frac{1}{2}(f(x^*) + f(y^*)) = f(x^*)$, ce qui contredit l'hypothèse que x^* minimise f sur Ω. □

Remarque 4. Cette proposition ne dit rien de l'existence d'un minimiseur.

Chapitre 2

Descente de gradient à pas optimal

On souhaite résoudre numériquement le problème de minimisation d'une fonction $f \in \mathcal{C}^0(\Omega)$ sur un ouvert $\Omega \subseteq \mathbb{R}^d$. Comme en général il n'est pas raisonnable d'espérer calculer de manière exacte un minimiseur ou même la valeur de l'infimum du problème (P), on cherchera à l'*approcher*. Il s'agira de construire une suite $(x^{(k)})_{k\geq 0}$ d'éléments de Ω vérifiant une des deux propriétés suivantes :

(a) la suite $x^{(k)}$ est minimisante pour (P), i.e. $\lim_{k\to+\infty} f(x^{(k)}) = \inf_\Omega f$.

(b) la suite $x^{(k)}$ converge vers un minimiseur de f sur Ω.

Dans ce chapitre, nous nous intéresserons à la construction de suites $x^{(k)}$ vérifiant la seconde propriété, qui est bien sûr plus forte que la première :.

2.1 Méthode de descente

Vocabulaire On appelle *méthode de descente* un procédé algorithmique permettant de construire itérativement une suite vérifiant (a) ou (b). Typiquement, une méthode de descente prend la forme suivante

$$\begin{cases} d^{(k)} = \dots & \textit{direction de descente} \\ t^{(k)} = \dots & \textit{pas de descente} \\ x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)} \end{cases}$$

Un tel algorithme est appelé méthode de descente si $f(x^{(k+1)}) \leq f(x^{(k)})$. Dans ce cours, on considèrera les possibilités suivantes :

(a) La *direction de descente* peut être égale à l'opposé du gradient, $d^{(k)} = -\nabla f(x^{(k)})$, auquel cas on parle de méthode de descente de gradient. Lorsque f est $\mathcal{C}^2$, il peut être plus avantageux de choisir $d^{(k)} = -\mathrm{D}^2 f(x^{(k)})^{-1} \nabla f(x^{(k)})$, auquel cas on parle de méthode de Newton.

(b) Le *pas de descente* peut être choisi constant ($t^{(k)} = \tau$), optimal (cf (2.2)), ou obtenu par des constructions un peu plus complexes, permettant de garantir la convergence de la méthode.

Définition 6. Soit $f \in \mathcal{C}^0(\Omega)$, $\Omega \subseteq \mathbb{R}^d$ ouvert. On appelle *direction de descente* en $x \in \Omega$ tout vecteur v tel que $\exists \tau > 0, \forall t \in [0, \tau], f(x + tv) < f(x)$.

Exercice 2. Si $f \in \mathcal{C}^1(\Omega)$ et $\langle v | \nabla f(x) \rangle < 0$, alors v est une direction de descente.

Si f est différentiable en x, on a $f(x+tv) = f(x) + \langle \nabla f(x) | v \rangle + \mathrm{o}(t)$. On cherche naturellement une direction de descente rendant le produit scalaire $\langle \nabla f(x) | v \rangle$ le plus petit possible, menant au choix de $d^{(k)} = -\nabla f(x^{(k)})$:

Lemme 10. *Soit $f \in \mathcal{C}^1(\Omega)$ et $x_0 \in \Omega$. Alors, $\min_{\|v\|=1} \mathrm{d}_{x_0} f(v) = -\|\nabla f(x_0)\|$ et si $\nabla f(x_0) \neq 0$, l'unique minimiseur est $v = -\nabla f(x_0) / \|\nabla f(x_0)\|$.*

Démonstration. On a $\mathrm{d}_x f(v) = \langle \nabla f(x_0) | v \rangle \geq -\|\nabla f(x_0)\| \|v\|$ par Cauchy-Schwarz, avec égalité si et seulement si v est positivement homogène à $-\nabla f(x_0)$. Comme $\|v\| = 1$, on a $v = -\nabla f(x_0) / \|\nabla f(x_0)\|$. □

2.2 Convexité et hessienne

Définition 7. Soit $f \in \mathcal{C}^2(\Omega)$, où $\Omega \subseteq \mathbb{R}^d$ est un ouvert. On appelle *hessienne* de f en $x_0 \in \Omega$ la matrice des dérivées partielles secondes :

$$\mathrm{D}^2 f(x) = \left(\frac{\partial^2 f}{\partial e_i \partial e_j}(x) \right)_{1 \leq i,j \leq d},$$

où l'on a noté $(e_i)_{1 \leq i \leq d}$ la base canonique de $\mathbb{R}^d$ et où

$$\frac{\partial^2 f}{\partial e_i \partial e_j}(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(\frac{\partial f}{\partial e_j}(x + \varepsilon e_i) - \frac{\partial f}{\partial e_j}(x) \right).$$

Définition 8. À toute matrice symétrique $A \in \mathcal{M}_{d,d}(\mathbb{R})$ on peut associer une fonction (appelée forme quadratique) $q_A : x \in \mathbb{R}^d \mapsto \langle x | Ax \rangle$.

(i) Une matrice symétrique A est dite *positive* (ce qu'on note $A \succeq 0$) si et seulement si la forme quadratique associée q_A est positive, i.e.

$$\forall x \in \mathbb{R}^d, \langle x | Ax \rangle \geq 0.$$

(ii) Une matrice symétrique est dite *définie positive* si

$$\forall x \in \mathbb{R}^d \setminus \{0\}, \langle x | Ax \rangle > 0.$$

(iii) Soient A, B deux matrices symétriques. On note $A \succeq B$ si et seulement si $A - B \succeq 0$, ou de manière équivalente si $q_A \geq q_B$.

Exemple 3. En particulier, on $\lambda \mathrm{Id} \preceq A \preceq \Lambda \mathrm{Id}$ si et seulement si

$$\forall x \in \mathbb{R}^d, \quad \lambda \|x\|^2 \leq \langle x | Ax \rangle \leq \lambda \|x\|^2.$$

Proposition 11. *Soit* $\Omega \subseteq \mathbb{R}^d$ *un ouvert,* $X \subseteq \Omega$ *convexe et* $f \in \mathcal{C}^2(\Omega)$*. Si* $\mathrm{D}^2 f(x) \succeq 0$*,* $\forall x \in X, \mathrm{D}^2 f(x) \succeq 0$*, alors* f *est convexe sur* X*.*

Remarque 5. La réciproque est fausse si X n'est pas ouvert.

Lemme 12. *Soit* $\Omega \subseteq \mathbb{R}^d$ *ouvert,* $f \in \mathcal{C}^2(\Omega)$*,* $x \in \Omega$ *et* $v \in \mathbb{R}^d$ *et* $g : t \mapsto f(x_t)$ *où* $x_t = x + tv$*. Alors,*

$$g''(t) = \langle \mathrm{D}^2 f(x_t) v | v \rangle. \tag{2.1}$$

Démonstration. Par le lemme 7, $g'(t) = \langle \nabla f(x_t) | v \rangle$, soit en coordonnées

$$g'(t) = \sum_{1 \le i \le d} \frac{\partial f}{\partial e_i}(x_t) v_i, = \sum_{1 \le i \le d} g_i(t) v_i,$$

où l'on a posé $f_i = \frac{\partial f}{\partial e_i}$ et $g_i(t) = f_i(x_t)$. Ainsi, $g''(t) = \sum_i g_i'(t) v_i$. En appliquant le lemme 7 à la fonction g_i on obtient

$$g_i'(t) = \langle \nabla f_i(x_t) | v \rangle = \sum_{1 \le j \le d} \frac{\partial g_i}{\partial e_j}(x_t) v_j = \sum_{1 \le j \le d} \frac{\partial^2 f}{\partial_j \partial e_i}(x_t) v_j,$$

de sorte que

$$g''(t) = \sum_{1 \le i,j \le d} \frac{\partial^2 f}{\partial e_j \partial e_i}(x_t) v_j v_i.$$

D'autre part, par définition de $\mathrm{D}^2 f$, les coordonnées du vecteur $\mathrm{D}^2 f(x_t) v$ sont

$$(\mathrm{D}^2 f(x_t) v)_i = \sum_{1 \le j \le d} \frac{\partial^2 f}{\partial e_i \partial e_j} v_j,$$

$$\text{d'où } \langle \mathrm{D}^2 f(x_t) v | v \rangle = \sum_{1 \le i,j \le d} \frac{\partial^2 f}{\partial e_i \partial e_j} v_j v_i = g''(x_t).$$ □

Démonstration de la proposition 11. Pour la réciproque, nous considérons $x, y \in \Omega$ et $g(t) = f(x_t)$ où $x_t = (1-t)x + ty$. Alors, $g''(t) = \langle \mathrm{D}^2 f(x_t)(y-x) | y - x \rangle$ est positif par hypothèse, de sorte que

$$g'(1) = g'(0) + \int_0^1 g''(t) \mathrm{d}t \ge g'(0),$$

soit $\langle \nabla f(y) | y - x \rangle \ge \langle \nabla f(x) | y - x \rangle$. La proposition 6 montre que f est convexe. □

2.3 Descente de gradient à pas optimal

Définition 9. Soit $\Omega \subseteq \mathbb{R}^d$ ouvert et $f \in \mathcal{C}^1(\Omega)$. L'algorithme de descente de gradient à pas optimal est donné par :

$$\begin{cases} d^{(k)} = -\nabla f(x^{(k)}) \\ t^{(k)} \in \arg\min_t f(x^{(k)} + td^{(k)}) \\ x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)} \end{cases} \tag{2.2}$$

Remarque 6. Par construction, les itérées vérifient $f(x^{(k+1)}) \leq f(x^{(k)})$.

Remarque 7. Pour pouvoir mettre en œuvre cet algorithme il faut pouvoir calculer le pas optimal $t^{(k)}$ à chaque itération, ce qui implique de résoudre un problème d'optimisation (sur $\mathbb{R}$). Ceci n'est faisable de manière exacte que dans un nombre très limité de cas. En général, on préfèrera d'autre méthode de calcul du pas.

Exemple 4. Soit $f : x \in \mathbb{R}^d \mapsto \frac{1}{2}\langle Qx|x\rangle + \langle b|x\rangle$ où Q est une matrice symétrique définie positive. Alors f est strictement convexe et $\nabla f(x) = Qx + b$. Soit $x^{(k)} \in \mathbb{R}^d$, $d^{(k)} = -\nabla f(x^{(k)})$. Pour calculer le pas $t^{(k)} \in \mathbb{R}$, on cherche le minimum de $g : t \in \mathbb{R} \mapsto f(x_t)$ où $x_t = x^{(k)} + td^{(k)}$. La fonction g est convexe et atteint donc son minimum en l'unique point $t^{(k)}$ vérifiant $g'(t^{(k)}) = 0$. Or, $g'(t) = \langle \nabla f(x_t)|d^{(k)}\rangle$, soit

$$\begin{aligned} g'(t^{(k)}) = 0 &\iff \langle Q(x^{(k)} + t^{(k)}d^{(k)}) + b|d^{(k)}\rangle = 0 \\ &\iff t^{(k)}\langle Qd^{(k)}|d^{(k)}\rangle - \langle d^{(k)}|d^{(k)}\rangle = 0. \end{aligned}$$

Ainsi, $t^{(k)} = \langle d^{(k)}|d^{(k)}\rangle / \langle Qd^{(k)}|d^{(k)}\rangle$. En résumé, dans le cas d'une fonction f de la forme considérée, l'algorithme de descente de gradient à pas optimal s'écrit

$$\begin{cases} d^{(k)} = -(Qx^{(k)} + b) \\ t^{(k)} = \frac{\langle d^{(k)}|d^{(k)}\rangle}{\langle Qd^{(k)}|d^{(k)}\rangle} \\ x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}. \end{cases} \tag{2.3}$$

On suppose dans la suite que l'ensemble

$$S = \{x \in \Omega \mid f(x) \leq f(x^{(0)})\} \text{ est compact,} \tag{2.4}$$

ce qui garantit (proposition 2) l'existence d'un minimiseur de f sur Ω.

Lemme 13. *Sous l'hypothèse* (2.4)*, le minimum dans la définition du pas est atteint.*

Démonstration. Il s'agit de démontrer que la fonction $g : t \mapsto f(x^{(k)} + td^{(k)})$ atteint son minimum, et grâce à la proposition 2, il suffit de montrer que le sous-niveau $S_g := \{t \in \mathbb{R} \mid g(t) \leq f(x^{(k)})\}$ est compact. Soit $(t_n)_{n\geq 1}$ une suite d'éléments de l'ensemble S_g. Alors $z_n = x^{(k)} + t_n d^{(k)}$ appartient au compact S (car par hypothèse $f(z_n) \leq f(x^{(k)}) \leq f(x^{(0)})$) et admet donc une sous-suite convergente. On en déduit que t_n possède aussi une sous-suite convergente, montrant la compacité de S_g. $\square$

Théorème 14. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe et $f \in \mathcal{C}^2(\Omega)$ vérifiant*
(i) le sous-niveau $S = \{x \in \Omega \mid f(x) \leq f(x^{(0)})\}$ est compact.
(ii) $\exists M > 0,\ \forall x \in S,\ \mathrm{D}^2 f(x) \preceq M\mathrm{Id}$,
(iii) f est strictement convexe,
Alors les itérées de l'algorithme (2.2) convergent vers l'unique minimiseur global de f sur Ω.

On utilisera la proposition suivante :

Proposition 15. *Soit $(x_k)_{k\geq 1}$ une suite bornée de $\mathbb{R}^d$ admettant une unique valeur d'adhérence $\bar{x}$.*[1] *Alors, $\lim_{k\to+\infty} x_k = \bar{x}$.*

Démonstration du théorème 14. Soit $k \geq 1$ et $g(t) = f(x^{(k)} + td^{(k)}) = f(x^{(k)} - t\nabla f(x^{(k)}))$. Par définition du pas optimal $t^{(k)}$ on a

$$f(x^{(k+1)}) = \min_t f(x^{(k)} + td^{(k)}) \leq \min_t g(t),$$

et nous allons utiliser un développement de Taylor pour estimer le minimum de g. Par le lemme 12, et en posant $x_t = x^{(k)} + td^{(k)}$ on a

$$g'(t) = \langle \nabla g(x_t) | d^{(k)} \rangle, \qquad g''(t) = \langle \mathrm{D}^2 g(x_t) d^{(k)} | d^{(k)} \rangle.$$

Comme sous-niveau d'une fonction convexe, l'ensemble S est un compact convexe, de sorte que l'ensemble $\sigma = \{t \in \mathbb{R} \mid x_t \in S\}$ est un segment. Par Taylor-Lagrange, pour tout $t \in S$, il existe $s \in [0,t]$ tel que $g(t) = g(0) + tg'(0) + \frac{t^2}{2} g''(s)$. Ainsi,

$$\begin{aligned} g(t) &= g(0) + tg'(0) + \frac{t^2}{2} g''(s) \\ &= f(x^{(k)}) - t \left\| \nabla f(x^{(k)}) \right\|^2 + \frac{t^2}{2} \langle \mathrm{D}^2 g(x_s) \nabla f(x^{(k)}) | \nabla f(x^{(k)}) \rangle \\ &\leq f(x^{(k)}) + \left(\frac{M}{2} t^2 - t \right) \left\| \nabla f(x^{(k)}) \right\|^2, \end{aligned}$$

où l'on a utilisé l'hypothèse (ii) pour passer de la deuxième à la troisième ligne. Le minimum de ce second membre est atteint en $t = 1/M$ et on a donc

$$f(x^{(k+1)}) \leq \min_t g(t) \leq f(x^{(k)}) - \frac{1}{2M} \left\| \nabla f(x^{(k)}) \right\|^2,$$

de sorte que

$$\left\| \nabla f(x^{(k)}) \right\|^2 \leq 2M(f(x^{(k)}) - f(x^{(k+1)})). \tag{2.5}$$

Ainsi, pour tout $K \geq 0$ on a

$$\sum_{0 \leq k \leq K} \left\| \nabla f(x^{(k)}) \right\|^2 \leq 2M(f(x^{(0)}) - f(x^{(K)})) \leq 2M(f(x^{(0)}) - \inf_\Omega f).$$

1. On rappelle que $\bar{x}$ est valeur d'adhérence de la suite $(x_k)_{k\geq 1}$ si et seulement si il existe une suite extraite $(x_{\sigma(k)})_{k\geq 1}$ dont la limite est $\bar{x}$

La série de terme général $\left\|\nabla f(x^{(k)})\right\|^2$ est donc convergente, d'où l'on déduit que $\lim_{k\to+\infty}\left\|\nabla f(x^{(k)})\right\| = 0$.

Montrons enfin que $\lim_{k\to+\infty} x^{(k)} = x^*$, où x^* est l'unique minimum de f sur Ω, l'unicité provenant de la stricte convexité de f et de la proposition 9. Comme $f(x^{(k)}) \leq f(x^{(0)})$, le point $x^{(k)}$ appartient à S, qui est par hypothèse compact donc borné. Pour montrer que la suite $x^{(k)}$ converge vers x^*, il suffit par la proposition 15 de démontrer qu'elle admet x^* pour seule valeur d'adhérence. Soit donc $(x^{(\sigma(k))})$ une sous-suite convergeant vers une valeur d'adhérence $\bar{x} \in S$. Alors, comme $f \in \mathcal{C}^1(\Omega)$, $\nabla f(\bar{x}) = \lim_{k\to+\infty} \nabla f(x^{(\sigma(k))}) = 0$. Par convexité (théorème 8), on sait que $\bar{x}$ est un minimiseur de f sur Ω. Par unicité du minimiseur, on en déduit que $\bar{x} = x^*$. $\square$

Chapitre 3

Descente de gradient à pas optimal II

3.1 Forte convexité et stabilité du minimum

Motivation. Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe $f \in \mathcal{C}^0(X)$ atteignant son minimum x^* sur Ω. On a vu dans la proposition 9 que si f est strictement convexe, alors f admet au plus un minimiseur x^* sur X. Une manière de reformuler cette propriété est la suivante :

$$(x \in \Omega \text{ et } f(x) \leq f(x^*)) \Longrightarrow x = x^*.$$

En pratique, nos algorithmes sont incapables de calculer x^* mais permettent au mieux de calculer une suite $x^{(k)}$ tel que $f(x^{(k)}) \leq f(x^*) + \varepsilon_k$ où $\lim_{k \to +\infty} \varepsilon_k = 0$. On aimerait pouvoir en déduire que $x^{(k)}$ est "proche" de la solution x^*, i.e. $\left\| x^{(k)} - x^* \right\| \leq C \varepsilon_k^\alpha$ pour un certain exposant $\alpha > 0$ et une constante $C > 0$. Nous verrons qu'une telle inégalité est vraie pour $\alpha = \frac{1}{2}$ si la fonction f est *fortement convexe*.

Définition 10. Soit $X \subseteq \mathbb{R}^d$ un ensemble convexe. Une fonction $f : X \to \mathbb{R}$ est dite *m-fortement convexe* sur X, où $m > 0$ si la fonction $f - \frac{m}{2} \|\cdot\|^2$ est convexe.

Exercice 3. Si f est m-fortement convexe, alors elle est aussi strictement convexe.

Proposition 16. *Soit Ω un ouvert de $\mathbb{R}^d$, $f \in \mathcal{C}^1(\Omega)$ et $X \subseteq \Omega$ convexe. Les propositions suivantes sont équivalentes :*

(i) f *est m-fortement convexe sur* X ;

(ii) $\forall x, y \in X, \quad f(y) \geq f(x) + \langle \nabla f(x) | y - x \rangle + \frac{m}{2} \|x - y\|^2$

(iii) $\forall x, y \in X, \quad \langle \nabla f(y) - \nabla f(x) | y - x \rangle \geq m \|x - y\|^2$

De plus, si $f \in \mathcal{C}^2(\Omega)$ et si $\forall x \in \Omega$, $\mathrm{D}^2 f(x) \succeq m\mathrm{Id}$, alors f est m-fortement convexe sur X.

Démonstration. La fonction f est m-fortement convexe si et seulement si la fonction $g = f - \frac{m}{2} \|\cdot\|^2$ est convexe. L'équivalence est s'obtient et utilisant les proposition 6 11 pour caractériser la convexité de g. $\square$

Exemple 5. La fonction $f : t \mapsto \exp(t)$ n'est pas fortement convexe sur $\mathbb{R}$ (car $f'' = \exp(t)$ n'est pas minoré sur $\mathbb{R}$), mais elle est m-fortement convexe sur tout segment $[a, b]$ pour $m = \min_{t\in[a,b]} \exp(t)$. De même, fonction $t \mapsto t^4$ n'est pas fortement convexe sur $\mathbb{R}$ mais est fortement convexe sur tout segment ne contenant pas l'origine.

Corollaire 17. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe, $f \in \mathcal{C}^1(\Omega)$ une fonction m-fortement convexe et x^* un minimiseur de f sur Ω. Alors, pour tout $x \in \Omega$,*
(i) $\|x - x^*\|^2 \leq \frac{2}{m}(f(x) - f(x^*))$
(ii) $\|x - x^*\| \leq \frac{1}{m} \|\nabla f(x)\|$
(iii) $f(x) - f(x^*) \leq \frac{1}{2m} \|\nabla f(x)\|^2$

Remarque 8. Le point (i) montre immédiatement que le minimiseur de f est unique : si x est un autre minimiseur, on a $f(x) = f(x^*)$, de sorte que par (i), $x = x^*$. Mais il faut surtout retenir que si x est "presqu'un minimiseur", au sens où $f(x) \leq f(x^*)+\varepsilon$, alors x est "proche" du minimiseur x^*, plus précisément $\|x - x^*\| \leq \sqrt{2\varepsilon/m}$. On tire des conclusions similaires du point (ii) : si la condition d'optimalité ($\nabla f(x^*) = 0$) est "presque vérifiée", i.e. si $\|\nabla f(x)\| \leq \varepsilon$ est petit, alors $\|x - x^*\| \leq \varepsilon/m$.

Exemple 6. Pour $f(t) = t^4$, on a $x^* = 0$ et l'inégalité (i), qui s'écrit $t^2 \leq \frac{2}{m}t^4$, n'est vraie pour aucun $m > 0$.

Démonstration. (i) est une conséquence immédiate de la formulation équivalente de la forte convexité donnée dans la proposition 16.(ii) :

$$f(x) \geq f(x^*) + \langle x - x^* | \nabla f(x^*)\rangle + \frac{m}{2} \|x - x^*\| = f(x^*) + \frac{m}{2} \|x - x^*\|^2 ,$$

où l'on a utilisé $\nabla f(x^*) = 0$ par optimalité de x^*. (ii) s'obtient de la même manière, en utilisant proposition 16.(iii).
(iii) Par l'inégalité (ii) de la proposition 16, on a pour tout $x, y \in \Omega$

$$f(y) \geq Q(x, y) := f(x) + \langle y - x|\nabla f(x)\rangle + \frac{m}{2} \|y - x\|^2 .$$

Ainsi, $f(y) \geq \inf_z g(z)$ où $g(z) = Q(x, z)$. Comme g est convexe sur $\mathbb{R}^d$, son unique minimiseur z^* est solution de l'équation

$$\nabla g(z^*) = 0 = \nabla f(x) + m(z^* - x),$$

c'est-à-dire que $z^* = x - \frac{1}{m}\nabla f(x)$. Ainsi, nous venons de démontrer que

$$\forall y \in \Omega, f(y) \geq g(z^*) = f(x) - \frac{1}{m} \|\nabla f(x)\|^2 + \frac{m}{2} \left\| \frac{1}{m}\nabla f(x) \right\|^2 = f(x) - \frac{1}{2m} \|\nabla f(x)\|^2 .$$

En prenant $y = x^*$ dans l'inégalité précédente, on obtient le résultat voulu. $\square$

3.2 Vitesse de convergence du gradient à pas optimal

Définition 11 (Vitesse de convergence)**.** Soit $(u_k)_{k\geq 0}$ une suite de limite u^*. On dira que la suite (u_k) converge *linéairement* vers u^* s'il existe $\kappa < 1$ et $k_0 \in \mathbb{N}$ telle que

$$\forall k \geq k_0, \|u_{k+1} - u^*\| \leq \kappa \|u_k - u^*\|.$$

On dira que la suite (u_k) converge *quadratiquement* vers u^* s'il existe $\gamma > 0$ telle que

$$\|u_{k+1} - u^*\| \leq \gamma \|u_k - u^*\|^2.$$

Théorème 18. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe et $f \in \mathcal{C}^2(\Omega)$ vérifiant*
(i) le sous-niveau $S = \{x \in \Omega \mid f(x) \leq f(x^{(0)})\}$ est compact.
(ii) $\exists M,\ \forall x \in S,\ \mathrm{D}^2 f(x) \preceq M\mathrm{Id}$,
(iii) $\exists m > 0,\ \forall x \in S,\ \mathrm{D}^2 f(x) \succeq m\mathrm{Id}$,
Alors les itérées de l'algorithme (2.2) *convergent vers l'unique minimiseur global de f sur Ω, et de plus, en posant $c = M/m$,*

$$f(x^{(k+1)}) - f(x^*) \leq \left(1 - \frac{1}{c}\right)\left(f(x^{(k)}) - f(x^*)\right). \tag{3.1}$$

Remarque 9. En d'autres termes, $(f(x^{(k)}))_{k\geq 0}$ converge linéairement vers $f(x^*)$.

Remarque 10. Si une fonction $f \in \mathcal{C}^2(\Omega)$, vérifie $m\mathrm{Id} \preceq \mathrm{D}^2 f(x) \leq M\mathrm{Id}$, on pourra dira que le *conditionnement* de f est majoré par $c = \frac{M}{m}$. L'inégalité (3.1) montre que cette quantité est cruciale pour comprendre la vitesse de convergence de l'algorithme de descente de gradient. Par exemple, si l'on souhaite estimer $f(x^*)$ à $\varepsilon > 0$ près, d'après l'inégalité (3.1), il suffit d'interrompre l'algorithme de descente de gradient à pas optimal après k itérations où

$$\left(1 - \frac{1}{c}\right)^k (f(x^{(0)}) - f(x^{(*)})) \leq \varepsilon,$$

soit

$$k \geq \frac{\log\left(\frac{f(x^{(0)}) - f(x^{(*)})}{\varepsilon}\right)}{\log\left(\frac{1}{1-1/c}\right)} \sim_{c\to+\infty} c\log\left(\frac{f(x^{(0)}) - f(x^{(*)})}{\varepsilon}\right)$$

Ainsi, dans le cas $c \gg 1$, le nombre d'itération est *proportionnel* au conditionnement !

Démonstration du théorème 18. Le corollaire (17) nous donne

$$f(x^k) - f(x^*) \leq \frac{1}{2m}\left\|\nabla f(x^{(k)})\right\|^2.$$

En combinant avec l'inégalité (2.5), on obtient

$$2m(f(x^k) - f(x^*)) \leq \left\|\nabla f(x^{(k)})\right\|^2 \leq 2M(f(x^{(k)}) - f(x^{(k+1)})),$$

de sorte qu'en posant $c = M/m$, on obtient bien l'inégalité voulue. □

3.3 Une condition suffisante pour la forte convexité

Proposition 19. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe, $f \in \mathcal{C}^2(\Omega)$ et $S \subseteq \Omega$ compact tel que $\forall x \in S$, $\mathrm{D}^2 f(x)$ est définie positive. Alors $\exists M \geq m > 0$ tels que*

$$\forall x \in S, m\mathrm{Id} \leq \mathrm{D}^2 f(x) \leq M\mathrm{Id}.$$

Démonstration. L'ensemble $K = \{(x,v) \in S \times \mathbb{R}^d \mid \|v\| = 1\}$ est compact comme fermé borné. La fonction la fonction $(x,v) \in K \times K \mapsto \langle \mathrm{D}^2 f(x) v | v\rangle$ est continue et donc bornée sur K, par exemple

$$\forall (x,v) \in K, m \leq \langle \mathrm{D}^2 f(x) v | v\rangle \leq M.$$

De plus la fonction atteint ses bornes sur K : il existe donc $(x_0, v_0) \in K$ tels que $m = \langle \mathrm{D}^2 f(x_0) v_0 | v_0\rangle > 0$ par hypothèses et comme $v_0 \neq 0$ (car $\|v_0\| = 1$). On en déduit l'inégalité voulue en remarquant que

$$\forall w \in \mathbb{R}^d, \langle \mathrm{D}^2 f(x) w | w\rangle = \langle \mathrm{D}^2 f(x) \frac{w}{\|w\|} | \frac{w}{\|w\|}\rangle \|w\|^2,$$

de sorte que $m \|w\|^2 \leq \langle \mathrm{D}^2 f(x) w | w\rangle \leq M \|w\|^2$ comme souhaité. □

Cette proposition permet d'énoncer une variante non-quantitative du théorème 18

Corollaire 20. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe et $f \in \mathcal{C}^2(\Omega)$ vérifiant*
(i) le sous-niveau $S = \{x \in \Omega \mid f(x) \leq f(x^{(0)})\}$ est compact,
(ii) $\forall x \in S$, $\mathrm{D}^2 f(x)$ est définie positive.
Alors les itérées de l'algorithme (2.2) *convergent vers l'unique minimiseur global x^* de f sur Ω, et de plus $(f(x^{(k)}))_{k \geq 0}$ converge linéairement vers $f(x^*)$.*

Exemple 7. Soit $a_1, \ldots, a_N$ une famille de vecteurs non nuls de $\mathbb{R}^d$ et

$$f(x) = -\sum_{1 \leq \alpha \leq N} \log(1 - \langle x | a_\alpha\rangle).$$

On fait les deux hypothèses [1] sur la famille $(a_\alpha)_{1 \leq \alpha \leq N}$, qui sont par exemple vérifiées par le choix de $a_1 = (1,0)$, $a_2 = (-1,0)$, $a_3 = (0,1)$, $a_4 = (0,-1)$ dans $\mathbb{R}^2$:

- la famille est positivement génératrice, i.e. tout point de $\mathbb{R}^d$ peut s'écrire sous la forme $x = \sum_\alpha \lambda_\alpha a_\alpha$, avec $\lambda_\alpha \geq 0$.
- l'ensemble $K = \{x \in \mathbb{R}^d \mid \forall \alpha \in \{1, \ldots, N\}, \langle x | a_\alpha\rangle \leq 1\}$ est compact.

1. La première hypothèse implique en fait la seconde. En effet, pour tout $v \in \{\pm e_i \mid 1 \leq i \leq d\}$, il existe $(\lambda_\alpha)_{1 \leq \alpha \leq N}$ positifs tels que $v = \sum_\alpha \lambda_\alpha a_\alpha$. Ainsi $\langle x | v\rangle = \sum_\alpha \alpha_\alpha \langle x | a_\alpha\rangle \leq \sum_i \lambda_\alpha$ pour tout $x \in K$. On en déduit que K est bornée pour la norme $\|\cdot\|_\infty$ (exercice) et donc compact.

Nous allons appliquer le corollaire précédente pour démontrer la convergence de l'algorithme de descente de gradient à pas optimal partant d'un point quelconque $x^{(0)} \in \Omega$. On remarque d'abord que la fonction f est de classe $\mathcal{C}^\infty$ sur son ensemble de définition $\Omega = \{x \in \mathbb{R}^d \mid \forall \alpha, \langle x | a_\alpha \rangle < 1\}$, qui est un ouvert convexe (exercice). Commençons par démontrer la propriété (ii). Soit $x \in \Omega$, $v \in \mathbb{R}^d$ et $g(t) = f(x_t)$ où $x_t = x + tv$. Alors,

$$g'(t) = \langle \nabla f(x_t) | v \rangle = \sum_{1 \leq \alpha \leq N} \frac{\langle a_\alpha | v \rangle}{1 - \langle x_t | a_\alpha \rangle}$$

$$g''(t) = \langle \mathrm{D}^2 f(x_t) v | v \rangle = \sum_{1 \leq \alpha \leq N} \frac{\langle a_\alpha | v \rangle^2}{(1 - \langle x_t | a_\alpha \rangle)^2}$$

Ainsi, pour tout $(x, v) \in \Omega \times \mathbb{R}^d$, $\langle \mathrm{D}^2 f(x) v | v \rangle \geq 0$ et la fonction f est donc bien convexe. Montrons maintenant que $\mathrm{D}^2 f(x)$ est définie positive. Soit $v \in \mathbb{R}^d$ tel que $\langle \mathrm{D}^2 f(x) v | v \rangle = 0$. Alors, $\forall \alpha$, $\langle a_\alpha | v \rangle = 0$ et comme la famille $(a_\alpha)_{\alpha \in A}$ est génératrice, on doit donc avoir $v = 0$.

Il nous reste à démontrer que le sous-niveau $S = \{x \in \Omega \mid f(x) \leq f(x^{(0)})\}$ est compact, ce qui se déduit de la continuité de f et de la compacité de K.

Chapitre 4

Descente avec pas d'Armijo

Une des objections à l'algorithme de descente de gradient à pas optimal est le calcul du pas demande de résoudre un problème d'optimisation 1D à chaque itération. Nous verrons dans ce chapitre une manière simple de choisir le pas qui donne les même garanties de convergence. Nous allons également un peu relaxer l'hypothèse que $d^{(k)} = -\nabla f(x^{(k)})$ afin de préparer le terrain à l'analyse de l'algorithme de Newton dans le chapitre suivant. Nous supposerons donc que la direction de descente est définie de la manière suivante, où $B^{(k)}$ est une matrice symétrique définie positive

$$d^{(k)} = -B^{(k)}\nabla f(x^{(k)}) \tag{4.1}$$

Le fait que $B^{(k)}$ est symétrique définie positive garantit que $\langle d^{(k)} | \nabla f(x^{(k)})\rangle < 0$, i.e. que $d^{(k)}$ est bien une direction de descente.

Remarque 11. Une méthode de descente où la direction $d^{(k)}$ est définie par la relation (4.1) est souvent appelé *descente de gradient préconditionné*, et la matrice $B^{(k)}$ est appelée *préconditionneur*.

4.1 Choix du pas par rebroussement

4.1.1 Rebroussement naïf

Comme nous souhaitons construire une méthode de descente, il serait assez naturel de considérer le pas suivant,

$$t^{(k)} = \max\{t \geq 0 \mid \exists \ell \in \mathbb{N}, t = 2^{-\ell} \text{ et } f(x^{(k)} + td^{(k)}) \leq f(x^{(k)})\}.$$

Autrement dit étant donné $x^{(k)}, d^{(k)}$ on teste les pas de $1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^\ell}$ et on s'arrête dès que que la condition de descente $f(x^{(k)} + 2^{-\ell}d^{(k)}) \leq f(x^{(k)})$ est satisfaite. Cet algorithme, assez naturel, est en fait non-convergent en général. Considérons $f(x) = x^2$ sur $\mathbb{R}$ et $x^{(0)} = 1$, $d^{(k)} = -f'(x^{(k)}) = -2x^{(k)}$. Alors le pas $t^{(k)} = 1$ est admissible et on obtient donc $x^{(k)} = (-1)^k$ qui ne converge pas vers le minimum de f sur Ω.

4.1.2 Rebroussement d'Armijo

L'idée est de renforcer la condition de descente $f(x^{(k)} + t^{(k)} d^{(k)}) \leq f(x^{(k)})$ par la condition plus forte (4.2) appellée "condition d'Armijo" :

Lemme 21. *Soit $f \in \mathcal{C}^1(\Omega)$, $\Omega \subseteq \mathbb{R}^d$ ouvert, $v \in \mathbb{R}^d$ tel que $\langle \nabla f(x) | v \rangle < 0$, et $0 < \alpha < \frac{1}{2}$. Alors, il existe $t_0 > 0$ tel que $\forall t \in [0, t_0]$,*

$$f(x + tv) \leq f(x) + \alpha t \langle \nabla f(x) | v \rangle. \tag{4.2}$$

Démonstration. C'est une conséquence immédiate de l'égalité

$$f(x + tv) = f(x) + \alpha t \langle \nabla f(x) | v \rangle + (1 - \alpha) t \langle \nabla f(x) | v \rangle + \mathrm{o}(t). \qquad \square$$

Définition 12 (Pas d'Armijo)**.** Soit $f \in \mathcal{C}^1(\Omega), \Omega$ ouvert convexe. Étant donné $x^{(k)}$, $d^{(k)}$ on appellera *pas d'Armijo* de paramètres $0 < \alpha < \frac{1}{2}$ et $0 < \beta < 1$ le pas $t^{(k)}$ défini par

$$\begin{aligned} t^{(k)} = \max\{t \mid \ell \in \mathbb{N}, t = \beta^\ell, x^{(k)} + t d^{(k)} \in \Omega \\ f(x^{(k)} + t d^{(k)}) \leq f(x^{(k)}) + t\alpha \langle \nabla f(x^{(k)}) | d^{(k)} \rangle\} \end{aligned} \tag{4.3}$$

Remarque 12. En d'autres termes, le pas d'Armijo peut être calculé par la procédure

`pas_armijo`$(x^{(k)}, d^{(k)})$:
  $t \leftarrow 1$
  tant que $f(x^{(k)} + t d^{(k)}) > f(x^{(k)}) + t\alpha \langle \nabla f(x^{(k)}) | d^{(k)} \rangle$:
    $t \leftarrow \beta t$
  retourner t

Pour écrire cet algorithme, on a supposé que lorsque x n'appartient pas au domaine de définition Ω de f alors $f(x) = +\infty$. Ceci qui garantit que si un pas t vérifie le critère d'Armijo

$$f(x^{(k)} + t d^{(k)}) \leq f(x^{(k)}) + t\alpha \langle \nabla f(x^{(k)}) | d^{(k)} \rangle$$

alors *automatiquement*, $x^{(k)} + t d^{(k)} \in \Omega$, et il n'est donc pas nécessaire de tester cette condition. Cette convention est très utile en pratique : il faudra toujour faire en sorte que la fonction Python évaluant la valeur $f(x)$ retourne `Inf` lorsque $x \notin \Omega$.

Définition 13 (Méthode de descente avec pas d'Armijo)**.** On appelle algorithme de descente de gradient préconditionné la méthode itérative suivante, où pour tout $k \in \mathbb{N}$, $B^{(k)} = A^{(k)} A^{(k)}$ et $A^{(k)}$ est une matrices symétriques définies positives :

$$\begin{cases} d^{(k)} = -B^{(k)} \nabla f(x^{(k)}) \\ t^{(k)} = \text{ donné par (4.3)} \\ x^{(k)} = x^{(k)} + t^{(k)} d^{(k)}. \end{cases}$$

4.1.3 Convergence de l'algorithme de descente de gradient préconditionné à rebroussement

Théorème 22. *Soit $\Omega \subseteq \mathbb{R}^d$ un ouvert convexe et $f \in \mathcal{C}^2(\Omega)$ vérifiant*
(i) le sous-niveau $S = \{x \in \Omega \mid f(x) \leq f(x^{(0)})\}$ est compact.
(ii) $\exists 0 < \lambda < \Lambda,\ \forall x \in S, \forall k \in \mathbb{N},\ \lambda \mathrm{Id} \leq A^{(k)} \mathrm{D}^2 f(x) A^{(k)} \preceq \Lambda \mathrm{Id}$,
Alors les itérées $x^{(k)}$ de l'algorithme de descente avec pas d'Armijo (Déf. 13) convergent vers l'unique minimiseur global de f sur Ω, et de plus, en posant $c = 2\alpha\lambda \min(1, \beta/\Lambda)$,

$$f(x^{(k+1)}) - f(x^*) \leq (1-c)\left(f(x^{(k)}) - f(x^*)\right). \tag{4.4}$$

Remarque 13. Cette analyse suggère une manière simple d'améliorer l'algorithme de descente de gradient : il suffit de choisir $B^{(k)}$ tel que le conditionnement Λ/λ soit aussi proche de 1 que possible. Dans le cas $f(x_1, x_2) = Kx_1^2 + x_2^2$, dont la hessienne est $H = \mathrm{diag}(K, 1)$ (constante), l'idéal est bien sûr de prendre $B = A^2$ où $A = \mathrm{diag}(K^{-1/2}, 1)$, soit $B = H^{-1}$. Plus généralement, le choix $B^{(k)} = \mathrm{D}^2 f(x^{(k)})^{-1}$ est souvent judicieux : il donne ce qu'on appelle la "méthode de Newton".

Lemme 23. *Soit $f \in \mathcal{C}^2(\Omega)$ et $B^{(k)} = (A^{(k)})^2$ telles que que*

$$\forall x \in S,\ A^{(k)} \mathrm{D}^2 f(x) A^{(k)} \leq \Lambda \mathrm{Id}.$$

Alors, le pas d'Armijo défini par (4.3) *vérifie*

$$t^{(k)} \geq \min\left(1, \frac{\beta}{\Lambda}\right).$$

Remarque 14. Cette inégalité permet de dire que l'algorithme `pas_armijo` calculant le pas d'Armijo termine dès que $\beta^k \leq \beta/\Lambda$. En d'autres termes, le nombre d'itération de l'algorithme de recherche du pas est au plus $\log(\Lambda/\beta)/\log(1/\beta)$.

Démonstration. Étant donné $t \in \mathbb{R}$, on pose $x_t = x^{(k)} + td^{(k)}$. Si t est tel que $x_t \in S$, alors le segment $[x_0, x_t]$ est inclus dans S et on montre donc comme dans la preuve du théorème 14 que

$$f(x_t) = f(x^{(k)}) + t\langle \nabla f(x^{(k)}) | d^{(k)} \rangle + \frac{t^2}{2} \langle D^2 f(x_s) d^{(k)} | d^{(k)} \rangle$$

Par (4.1), on a $d^{(k)} = -B^{(k)} \nabla f(x^{(k)}) = -A^2 \nabla f(x^{(k)})$ où on a noté $A = A^{(k)}$. Ainsi,

$$\begin{aligned}
\langle D^2 f(x_s) d^{(k)} | d^{(k)} \rangle &= \langle D^2 f(x_s) A^2 \nabla f(x^{(k)}) | A^2 \nabla f(x^{(k)}) \rangle \\
&= \langle A D^2 f(x_s) A (A \nabla f(x^{(k)})) | A \nabla f(x^{(k)}) \rangle \\
&\leq \Lambda \left\| A \nabla f(x^{(k)}) \right\|^2 = \Lambda \langle A^2 \nabla f(x^{(k)}) | \nabla f(x^{(k)}) \rangle = -\langle \nabla f(x^{(k)}) | d^{(k)} \rangle
\end{aligned}$$

de sorte que

$$f(x_t) \leq f(x^{(k)}) + (1 - \frac{\Lambda}{2}t)t\langle \nabla f(x^{(k)}) | d^{(k)} \rangle.$$

Ainsi, la condition d'Armijo

$$f(x_t) \leq f(x^{(k)}) + \alpha t \langle \nabla f(x^{(k)}) | d^{(k)} \rangle$$

est vérifiée dès que $1 - \frac{\Lambda}{2}t \geq \alpha$. Comme $\alpha \leq \frac{1}{2}$, il suffit pour cela que $t \leq \Lambda$. □

Lemme 24. *Soit $X \subseteq \Omega \subseteq \mathbb{R}^d$ avec X convexe et Ω ouvert. Soit $f \in \mathcal{C}^2(\Omega)$ et A une matrice symétrique vérifiant $\forall x \in X, A\mathrm{D}^2 f(x) A \geq \lambda \mathrm{Id}$. Alors*

$$2\lambda(f(x) - f(x^*)) \leq \|A\nabla f(x)\|^2 .$$

Démonstration. On considère $g(x) = f(Ax)$. Alors $\nabla g(x) = A\nabla f(x)$ et $\mathrm{D}^2 g(x) = A\mathrm{D}^2 f(x) A$. On a supposé $A\mathrm{D}^2 f(x) A \geq \lambda \mathrm{Id}$, et par le corollaire (17) on a

$$2\lambda(g(y) - g(y^*)) \leq \|\nabla g(y)\|^2 ,$$

où $y^* = Ax^*$. Ainsi, en posant $y = Ax$, $2\lambda(f(x) - f(x^*)) \leq \|A\nabla f(x)\|^2$. □

Démonstration du théorème 22. Par définition du pas d'Armijo, on sait que

$$\begin{aligned}
f(x^{(k+1)}) &\leq f(x^{(k)}) + \alpha t^{(k)} \langle \nabla f(x^{(k)}) | d^{(k)} \rangle \\
&= f(x^{(k)}) - \alpha t^{(k)} \langle \nabla f(x^{(k)}) | B^{(k)} \nabla f(x^{(k)}) \rangle \\
&= f(x^{(k)}) - \alpha t^{(k)} \langle A^{(k)} \nabla f(x^{(k)}) | A^{(k)} \nabla f(x^{(k)}) \rangle \\
&\leq f(x^{(k)}) - \varepsilon \left\| A^{(k)} \nabla f(x^{(k)}) \right\|^2 \quad \text{où } \varepsilon = \alpha \min(1, \beta/\Lambda)
\end{aligned}$$

où on a utilisé le lemme 23 pour minorer $t^{(k)}$. En combinant avec le lemme précédent on obtient

$$\begin{aligned}
&f(x^{(k+1)}) \leq f(x^{(k)}) - 2\lambda\varepsilon(f(x^k) - f(x^*)) \\
&f(x^{(k+1)}) - f(x^*) \leq (1 - c)(f(x^{(k)}) - f(x^*)) \text{ où } c = 2\lambda\varepsilon.
\end{aligned}$$

□